Staging Between Pleiades and AWS within PBS Jobs Submitted from a PFE

This article describes how to use `stagein` and `stageout` directives in a PBS job script to copy files between Pleiades and the $PBS_O_WORKDIR directory in AWS.

When you submit a job from Pleiades, the value of the job's $PBS_O_WORKDIR variable (the `cwd` at the time of the `qsub` operation) is recreated in AWS. If the job uses your persistent /nobackup directory, it will be under that filesystem. Otherwise, the value will be a symbolic link created under the node's root filesystem that points to the primary job-time filesystem. This enables the job to change (`cd`) to the $PBS_O_WORKDIR directory.

For jobs submitted from Pleiades, the $PBS_O_WORKDIR is restricted to one under your /nobackup directory, since you can only submit batch jobs to AWS from a Pleiades /nobackup directory.

```
#CLOUD -stagein_file=path_to_file/file
#CLOUD -stagein_dir=path_to_dir
#CLOUD -stageout_file=path_to_file/file
#CLOUD -stageout_dir=path_to_dir
#CLOUD -stageout_file_delete=path_to_file/file
#CLOUD -stageout_dir_delete=path_to_dir
```

Note: In a PBS script, all PBS directives should be specified before the CLOUD directives.

More details about staging are described below:

Using the stagein Directive

You can `stagein` files or directories under your Pleiades /nobackup directory, but not under your Pleiades /home directory.

All Pleiades files to be used in AWS must be specified in the PBS script. You can give individual files (wildcards included) or individual directories.

TIP: If you have a lot of files to stage in, you can tar up the directory, and just stage in the tar file. This will speed up the process. (You will need to untar the file inside your PBS script.)
You can use either an absolute or a relative path. When a relative path is used, it is relative to the $PBS_O_WORKDIR directory on Pleiades where the `qsub` command was issued.

The staged-in files and/or directories will have the same directory structure on AWS as they have on Pleiades. In the two examples below, the PBS job is submitted from `/nobackup/username/run` on Pleiades.

Example 1

The following directive will result in `/nobackup/username/run/input/file` on Pleiades to be staged into `/nobackup/username/run/input/file` on AWS, *not* `/nobackup/username/run/file` on AWS:

```
#CLOUD -stagein_file=input/file
```

Example 2

The following directive will copy **/nobackup/username/src/mpi_pi.f** from Pleiades to **/nobackup/username/src/mpi_pi.f** on AWS:

```
#CLOUD -stagein_file=/nobackup/username/src/mpi_pi.f
```

Using the stageout Directive

Only the relative path should be used. In addition, staging out files to Pleiades at a relative path above $PBS_O_WORKDIR (such as ../src) is not supported. The staged out files and/or directories will have the same directory structure on Pleiades as they have on AWS.

Stageout and Delete

Optionally, you can use the following directives to perform two tasksâ   stage out, then delete:

```
#CLOUD -stageout_file_delete=output
#CLOUD -stageout_dir_delete=test_dir
```

Order of Operation

- **-stagein_file** and **-stagein_dir** operations occur in the PBS prologue.
- **-stageout_file** and **-stageout_dir** operations occur at the end of job, specifically:
    - For AWS filesystems with a distinct filesystem server, such as a persistent filesystem or a shared-type job-time filesystem, the operations occur on the filesystem server after the job exits.
    - For other types of job-time filesystems, the operations occur in the PBS epilogue.

If you are using other **CLOUD** directives, see Order of Operation of Cloud Directives.